Comparison of Tonality Models in Measuring Chord Sequence Similarity

Saebyul Park, Jeounghoon Kim, Juhan Nam
GSCT, KAIST, Republic of Korea
saebyul_park@kaist.ac.kr, miru@kaist.ac.kr, juhannam@kaist.ac.kr

ABSTRACT

The objective of this study is to examine chord sequence similarity measures and experimentally assess their relationship with human perception of the similarity. We present five different types of chord sequence similarity measures based on different tonality models, including 1) tonal dissonance of intervals, 2) circle of fifths, 3) harmonic relations, 4) tonal pitch space, and 5) hierarchy of harmonic stability. For the evaluation, we collected 50 chord sequence pairs from US musical copyright infringement cases. Also, we surveyed human evaluation to compare it with the computational chord sequence similarity. The results show that those based on tonal pitch space and hierarchy of harmonic stability are relatively more correlated with the human judgement data and a combination of the two similarity measures further increases the correlation.

I. INTRODUCTION

Measuring music similarity has always received high attention in the fields of music information retrieval (MIR) and music psychology research, since classification and analysis of music basically start by concentrating on how individual songs share similar musical characteristics. Especially, measuring music similarity based on *harmony*, which is one of the core features of western tonal music, has great advantages in plagiarism detection, genre classification, identifying cover song and music recommendation systems (Casey et al., 2008).

Recently, various research on harmonic similarity has been carried out in the form of comparing symbolic chord sequences (B. De Haas, Veltkamp, & Wiering, 2008; W. B. De Haas, Wiering, & Veltkamp, 2013; Freedman, 2015; Hanna, Robine, & Rocher, 2009; Rocher, Robine, Hanna, & Desainte-Catherine, 2010). Even though these studies made great contributions in systematically and quantitatively measuring chord sequence similarity encompassing music theoretical models, we believe that there are additional musical factors that deserve to be included in the analysis. Moreover, previous work rarely has considered the relationship between computational measures of chord sequence similarity and human perception of the similarity.

In this paper, we present five different types of chord sequence similarity measures based on different tonality models and compare them to human judgements of the similarity. Through the experiment, we evaluate the computational similarity measures and show how they are correlated to the human data.

II. METHODS

A. Study 1: Computational Chord Sequence Similarity

For the experiment, we collected 50 chord sequence pairs extracted from songs in US musical copyright infringement cases as they were regarded as similar, but not identical. Specifically, we normalized the key by transposing all songs to C major. We cut the chord sequences into a set of 4 chords from each of song pairs and limited the chords to triads in major and minor mode (i.e., major triad, minor triad, diminished triad, and augmented triad). In order to compute chord similarity measures, we represented the chords as a triplet of pitch classes in a range of 0 to11. For example, C major triad is represented as “0 4 7”, C minor triad is represented as “0 3 7” and A minor triad is represented as “9 0 4”.

The computational similarity measures are based on edit distance in common. The edit distance, also known as the Levenshtein algorithm, is a metric that computes the minimum number of operations needed to transform one sequence into the other. The operations between sequences include deletion, insertion, and substitution of symbols. Let y be an operation cost function, e the empty string, two chord sequences, $a = a_1 \ldots a_n$ and $b = b_1 \ldots b_n$. Then, the edit distance matrix, d_{mn} between a and b is computed as follows:

$$d_{00} = 0,$$
$$d_{i0} = i \quad for\ 1 \le i \le m,$$
$$d_{0j} = j \quad for\ 1 \le j \le n$$
$$d_{ij} = \begin{cases} d_{i-1,j-1} & for\ a_j = b_i \\ min \begin{cases} d_{i-1,j} + y_i(e \to b_j) \\ d_{i,j-1} + y_d(a_i \to e) \\ d_{i-1,j-1} + y_s(a_j \to b_i) \end{cases} & for\ a_j \neq b_i \end{cases}$$
$$for\ 1 \le i \le m, 1 \le j \le n$$

While fixing both insertion and deletion costs to a constant value of 1, we vary the substitution cost based on distance between the two chords derived from different tonality models. In the following subsections, we describe how to quantify the substitution cost in each of the models.

1) *Tonal dissonance of intervals.* With this model, we use the root of the triads only and thus measure the chord distance (i.e., the substitution cost) from the two root notes. Table 1 shows the dissonance rating of interval (Nordmark & Fahlén, 1988). We compute the substitution cost by taking the number of semitones from 1 to 11 and normalizing the dissonance rating values to the range between 1 and 2.

2) *Circle of fifths* The second substitution cost is based on the circle of fifths, which is a graphical representation of the 12 notes scale placed onto a circle where neighbouring notes are separated by a fifth interval (Figure 1). With this model, we also use the root of the triads only. The chord distance is defined as the smaller number of steps in the circle of fifths either clockwise or counter-clockwise. We compute the substitution cost by normalizing the number of non-zero steps to the range between 1 and 2.

Table 1. Dissonance rating between two intervals (Cook, 1999; Nordmark & Fahlén, 1988)

Interval Name	Number of Semitones	Dissonance Rating(1-7)
Octave	12	1.7
Fifth	7	1.7
Fourth	5	2.0
Major third	4	2.0
Major sixth	9	2.4
Minor third	3	2.6
Minor sixth	8	3.0
Minor seventh	10	3.3
Major second	2	3.9
Tritone	6	4.0
Major seventh	11	5.3
Minor second	1	5.7
Minor ninth	13	5.8

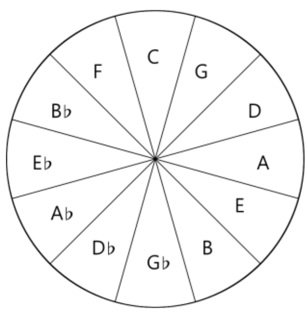

Figure 1. The Circle of fifths

3) *Harmonic relations* The third substitution cost function is defined based on the theoretical harmonic relations (Lerdahl, 2001). The distance between chord pairs, which all belong to the harmonic function in C major as we transposed all chords to this key, are selected from the theoretical harmonic relations (See Table 2). The values between 5 and 8 are scaled to the range between 1 and 2 as a substitution cost.

Table 2. Theoretical Harmonic Relations from Lerdahl (2001)

	Ⅰ	Ⅱ	Ⅲ	Ⅳ	Ⅴ	Ⅵ	Ⅶ
Ⅰ	0						
Ⅱ	8	0					
Ⅲ	7	8	0				
Ⅳ	5	7	8	0			
Ⅴ	5	5	7	8	0		
Ⅵ	7	5	5	7	8	0	
Ⅶ	8	7	5	5	7	8	0

4) *Tonal pitch space* For the fourth substitution cost function, the tonal pitch space (TPS) model (Lerdahl, 2001) was used to calculate a distance between two chords. The basis of the model is the basic space (see Figure 2) which comprises five hierarchical levels (a-e) consisting of pitch class subsets ordered from root level to chromatic level. The chordal distance is calculated by the number of non-common pitch classes within the basic spaces of two chords, divided by two. We calculated the distance from level a to c (level a-c), as we limited our chord selection to triads only, resulting the maximum distance of 6. Again, the non-zero chord distance is scaled to the range between 1 and 2 as a substitution cost.

(a) root level:	0											
(b) fifth level:	0							7				
(c) triadic level:	0				4			7				
(d) diatonic level:	0		2		4	5		7		9		11
(e) chromatic level	0	1	2	3	4	5	6	7	8	9	10	11

Figure 2. Diatonic basic space for C major triad (C=0, C#=1…B=11)

5) *Hierarchy of harmonic stability* For the fifth substitution cost function, we use the distance principle derived from hierarchy of harmonic stability are applied (Bharucha & Krumhansl, 1983). In this model, the psychological distances between chords reflect both key membership and stability within the key: 1) Chords from the same key are perceived as more closely related than those from different keys, 2) Chords in a harmonic core (Tonic, subdominant and dominant) are perceived as more closely related to each other compared to other chords from the key, but not in core. Based on this model, the substitution cost is computed when two different chords C_1 and C_2 are given as followed:

$$\begin{array}{lll} cost = 0 & \text{when} & C_1 = C_2 \\ cost = 1 & \text{when} & C_1, C_2 \in K \text{ and } C_1, C_2 \in S \\ cost = 2 & \text{when} & C_1, C_2 \in K \text{ and } C_1 \text{ or } C_2 \notin S \\ cost = 3 & \text{when} & C_1 \text{ or } C_2 \notin K, \end{array}$$

where K contains the seven triads built upon the seven degrees of the diatonic scale and S is the set containing the three harmonic chords (I, IV, and V) of the harmonic core. The non-zero substitution cost is scaled to the range between 1 and 2 to be used in the edit distance.

B. Study 2: Human Survey Experiment

The objective of the survey experiment is to gather human similarity judgment data to compare the results with the computed similarity values. A total of 28 subjects participated in the experiment. Participants were 24–58 years old (mean μ = 27.7) and were recruited without regard to their musical training.

Chord sequence pairs generated using US copyright infringement cases from study 1 were also used in audio form in study 2. Chords were generated in a C major key, within a range of two octaves centred at middle C. Chord sequences were composed of 4 chords, with a beat of a quarter notes each (140 bpm), resulting the total length of the audio of 6 seconds.

Human experiments were conducted as an on-line survey. Before starting to answer the questionnaire, participants were informed about the task and presented with examples of very similar and very dissimilar chord sequences. During the survey, participants listened to the audio clips for each chord sequence pair, and were asked to rate the similarity of these pairs on a scale from 1 to 4 (1=very similar, 2=similar, 3=dissimilar, and 4=very dissimilar). It took approximately 30 minutes to finish.

III. RESULTS

The comparison results are summarized in Figure 3. Pearson's correlation coefficient was used to compare the results between the computational measures and human survey data.

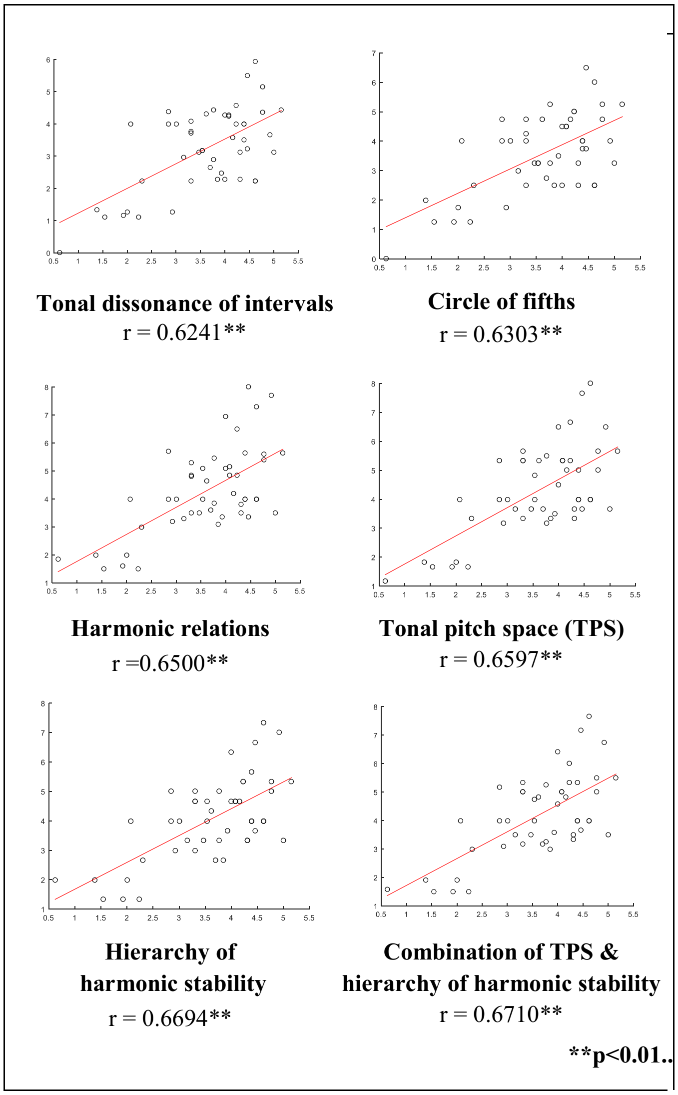

Figure 3. Pearson's correlation coefficient between computational measures and human judgements in chord sequence similarity

As shown in Figure 3, all of computed similarity measurements had significant correlation with human judgment of similarity ($p<0.01$). The highest correlation was achieved by the similarity measure based on Bharucha–Krumhansl's hierarchy of harmonic stability model ($r=0.669$, $p<0.01$), followed by the TPS model ($r=0.659$, $p<0.01$). This suggests that key membership and harmonic stability play an important role in the perception of similarity based on harmony, as well as the common pitch classes shared by two chords. Every correlation of measurements on chord level was higher than those using root notes only, which implies that more complex chord representations can enhance the performance of harmonic similarity measurements. A combination of results from the two highest-ranking models, hierarchy of harmonic stability and TPS, provides even higher correlation with the human judgements. This suggests that considering multiple tonality models can yield to better results than relying on a single model only.

IV. CONCLUSIONS

We presented five different types of chord sequence similarity measures based on different psychological distances of harmony. We also gathered human judgements of similarity between two chord sequence pairs. The main contribution of this study is to provide theoretical and empirical evidence on the relationship between computational approaches and human judgements on harmonic similarity. Furthermore, the results show that using a combination of similarity measures from different tonality models can improve correlation with human survey data. However, there are still numerous issues remaining to be addressed for the measurement of chord sequence similarity. In this study, we handled the substitution cost only while fixing the deletion and insertion costs to a constant value. Considering all cases separately and in conjunction will help comparing more diverse chord sequences and the importance of each operation. Finally, future work needs to take contextual information within a chord sequence into account.

REFERENCES

Bharucha, J., & Krumhansl, C. L. (1983). The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition, 13*(1), 63-102.

Casey, M. A., Veltkamp, R., Goto, M., Leman, M., Rhodes, C., & Slaney, M. (2008). Content-based music information retrieval: Current directions and future challenges. *Proceedings of the IEEE, 96*(4), 668-696.

Cook, P. R. (1999). *Music, cognition, and computerized sound: An introduction to psychoacoustics*: The MIT press.

De Haas, B., Veltkamp, R. C., & Wiering, F. (2008). *Tonal Pitch Step Distance: a Similarity Measure for Chord Progressions.* Paper presented at the ISMIR.

De Haas, W. B., Wiering, F., & Veltkamp, R. C. (2013). A geometrical distance measure for determining the similarity of musical harmony. *International Journal of Multimedia Information Retrieval, 2*(3), 189-202.

Freedman, D. (2015). *Correlational Harmonic Metrics: Bridging Computational and Human Notions of Musical Harmony.*

Hanna, P., Robine, M., & Rocher, T. (2009). *An alignment based system for chord sequence retrieval.* Paper presented at the Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries.

Lerdahl, F. (2001). *Tonal pitch space*: Oxford University Press.

Nordmark, J., & Fahlén, L. E. (1988). Beat theories of musical consonance. *KTH Royal Institute of Technology STL-QPSR, 29*(1), 111-122.

Rocher, T., Robine, M., Hanna, P., & Desainte-Catherine, M. (2010). *A survey of chord distances with comparison for chord analysis*: Ann Arbor, MI: Michigan Publishing, University of Michigan Library.